

### **REMARKS**

In the non-final Office Action, the Examiner rejected claims 18 and 19 under 35 U.S.C. § 112, first paragraph, as failing to comply with the written description requirement; and rejected claims 1-11 and 16-19 under 35 U.S.C. § 103(a) as unpatentable over Liddy et al. (U.S. Patent No. 6,026,388) in view of Grefenstette et al. (U.S. Patent No. 6,778,979) and Olsson et al. (U.S. Patent No. 6,266,667). Applicants traverse these rejections.<sup>1</sup>

By way of the present amendment, Applicants amend claims 18 and 19 to improve form. No new matter has been added by way of the present amendment. Claims 1-19 remain pending, of which claims 12-15 have been withdrawn from consideration by the Examiner.

Claims 18 and 19 stand rejected under 35 U.S.C. § 112, first paragraph, as allegedly failing to comply with the written description requirement. Specifically, the Examiner alleges that the specification does not disclose the features “predetermined number” and “predetermined number of the documents is equal to approximately half of the documents” in claims 18 and 19, respectively. Without acquiescing in the Examiner’s rejection, but merely to expedite prosecution, Applicants amend claims 18 and 19 to remove the word “predetermined.” As such, withdrawal of the rejection of claims 18 and 19 under 35 U.S.C. § 112, first paragraph, is respectfully requested.

Claims 1-11 and 16-19 stand rejected under 35 U.S.C. § 103(a) as allegedly unpatentable over Liddy et al. in view of Grefenstette et al. and Olsson et al. Applicants respectfully traverse this rejection.

---

<sup>1</sup> As Applicants’ remarks with respect to the Examiner’s rejections are sufficient to overcome these rejections, Applicants’ silence as to assertions by the Examiner in the Office Action or certain requirements that may be applicable to such rejections (e.g., whether a reference constitutes prior art, motivation to combine reference, assertions as to dependent claims, etc.) is not a concession by Applicants that such assertions are accurate or such requirements have been met, and Applicants reserve the right to analyze and dispute such assertions/requirements in the future.

Independent claim 1 is directed to a method of creating labels for clusters of documents. The method comprises identifying topics associated with the documents in the clusters; determining whether the topics are associated with at least half of the documents in the clusters; adding ones of the topics that are associated with at least half of the documents in the clusters to cluster lists; and forming labels for the clusters from the cluster lists. Liddy et al., Grefenstette et al., and Olsson et al., whether taken alone or in any reasonable combination, do not disclose or suggest this combination of features.

For example, Liddy et al., Grefenstette et al., and Olsson et al. do not disclose adding ones of topics that are associated with at least half of the documents in clusters to cluster lists, as required by claim 1. In rejecting claim 1, the Examiner admits that Liddy et al. does not disclose this feature (Office Action, pg. 4) and relies on column 49, lines 18-37 of Grefenstette et al. and column 2, lines 49-62 of Olsson et al. as allegedly disclosing this feature (Office Action, pp. 4-5). Applicants disagree.

At column 49, lines 18-37, Grefenstette et al. discloses:

In generating the set of categories 3620, the categorizer 3610 classifies input document to generate classification labels for the document content 3612. Terms and entities (i.e., typed terms, such as people organizations, locations, etc.) are extracted from the document content. For example, given a classification scheme such as a class hierarchy (e.g., from a DMOZ ontology that is available on the Internet at [dmoz.org](http://dmoz.org)) in which documents are assigned class labels (or assigned to nodes in a labeled hierarchy), a classification profile is derived that allows document content to be assigned to an existing label or to an existing class, by measuring the similarity between the new document and the known class profiles.

Document classification labels define the set of categories 3620 output by the categorizer 3610. These classification labels in one embodiment are appended to the query 3812 by query generator 3810 to restrict the scope of the query (i.e., the entity 3808 and the context vector 3822) to folders corresponding to classification labels in a document collection of an information retrieval system.

In this section, Grefenstette et al. discloses that a classification profile may be derived that allows document content to be assigned to an existing label or to an existing class by measuring the similarity between a new document and known class profiles. This section of Grefenstette et al. further discloses that classification labels may be appended to a query to restrict the scope of the query. This section of Grefenstette et al. discloses labeling document content and does not disclose or suggest adding ones of topics that are associated with at least half of the documents in clusters to cluster lists, as recited in claim 1. In fact, this section of Grefenstette et al. discloses that the qualifying factor for assigning content to labels is a measurement of the similarity between a new document and a known class profile. Clearly, this does not equate to adding ones of topics that are associated with at least half of the documents in clusters to cluster lists, as recited in claim 1.

At column 2, lines 49-62, Olsson et al. discloses:

Firefly, <http://www.firefly.net>, is one of the best-known social filtering systems. The technique is called Feature-Guided Automated Collaborative Filtering. This filtering technique builds a profile of each user with their opinions for different documents. The documents are divided in different groups (classes of documents) and for each group the users are clustered in a nearest "neighbor" style. For each group of documents the users' opinions in the same cluster are compared. To find documents to recommend, firefly matches all users in a cluster (this is somewhat simplified to increase understandability). If two users have approximately the same opinions for most of the documents in a group, but they have not read all of them, then it is likely that the users would like the unread documents of the group too.

This section of Olsson et al. discloses a social filtering system in which users' opinions of a cluster of documents are compared. If two users have approximately the same opinions for most of the documents in a group, but have not read all of them, then it is likely that the users would like the unread documents of the group, too. This section of Olsson et al. discloses comparing users' opinions of documents and does not disclose or suggest adding ones of topics that are associated with at least half of the documents in clusters to cluster lists, as recited in claim 1. The Examiner

appears to be alleging that Olsson et al.'s disclosure relating to recommending documents to users when they share a same opinion for "most documents in a group" corresponds to adding ones of topics that are associated with at least half of the documents in clusters to cluster lists, as recited in claim 1. Applicants disagree. Olsson et al.'s recommending documents in a group of documents, does not disclose or suggest adding ones of topics that are associated with at least half of the documents in clusters to cluster lists, as recited in claim 1.

For at least the foregoing reasons, Applicants submit that claim 1 is patentable over Liddy et al., Grefenstette et al., and Olsson et al., whether taken alone or in any reasonable combination.

Claims 2-8 depend from claim 1 and are, therefore, patentable over Liddy et al., Grefenstette et al., and Olsson et al., whether taken alone or in any reasonable combination, for at least the reasons given with regard to claim 1. Moreover, these claims recite additional features not disclosed or suggested by Liddy et al., Grefenstette et al., and Olsson et al.

For example, claim 5 recites ranking the ones of the topics based on a number of the documents with which the ones of the topics are associated. The Examiner relies on column 24, line 56 – column 25, line 2 and column 25, lines 30-37 of Liddy et al. and column 2, lines 49-62 of Olsson et al. as allegedly disclosing this feature (Office Action, pg. 6). Applicants respectfully disagree.

At column 24, line 56 – column 25, line 2, Liddy et al. discloses:

... the system will retrieve all documents with a non-zero score. Note that documents from different sources are interfiled and ranked in a single list.

The RP filtering function is accomplished by means of a multiple regression formula that successfully predicts cut-off criteria on a ranked list of relevant documents for individual queries based on the similarity of documents to queries as indicated by the vector matching (and optionally

the proper noun matching) scores. The RP is sensitive to the varied distributions of similarity scores (or match scores) for different queries, and is able to present to the user a certain limited percentage of the upper range of scored documents with a high probability that close to 100% recall will be achieved.

This section of Liddy et al. discloses predicting cut-off criteria on a ranked list of relevant documents for individual queries based on the similarity of documents to queries as indicated by vector matching scores. This section of Liddy et al. does not mention or relate to “a number of documents with which topics are associated,” as recited in claim 5. Therefore, this section of Liddy et al. cannot disclose or suggest ranking the ones of the topics based on a number of the documents with which the ones of the topics are associated, as recited in claim 5.

At column 25, lines 30-37, Liddy et al. discloses:

Headlines from newspaper articles or titles from documents in the cluster are used to form labels for clusters. Headlines or titles are selected from documents that are near the centroid of a particular cluster, and are therefore highly representative of the cluster's document contents. An alternative labeling scheme, selectable by the user, is the use of the labeled subject codes which make up either the centroid document's vector or the cluster vector.

This section of Liddy et al. discloses that headlines from newspaper articles or titles from documents that are near the centroid of a particular cluster are used to form labels for clusters. This section of Liddy et al. does not mention a number of documents with which topics are associated. Therefore, this section of Liddy et al. cannot disclose or suggest ranking the ones of the topics based on a number of the documents with which the ones of the topics are associated, as recited in claim 5.

As noted above, at column 2, lines 49-62, Olsson et al. discloses a social filtering system in which users' opinions of a cluster of documents are compared. If two users have approximately the same opinions for most of the documents in a group, but have not read all of them, then it is likely that the users would like the unread documents of the group, too. This section of Olsson et al. does

not disclose or suggest ranking the ones of the topics based on a number of the documents with which the ones of the topics are associated, as recited in claim 5.

For at least these additional reasons, Applicants submit that claim 5 is patentable over Liddy et al., Grefenstette et al., and Olsson et al., whether taken alone or in any reasonable combination.

Independent claims 9, 16, and 18 recite features similar to (but possibly different in scope from) features recited in claim 1. Claims 9, 16, and 18 are, therefore, patentable over Liddy et al., Grefenstette et al., and Olsson et al., whether taken alone or in any reasonable combination, for at least reasons similar to reasons given with regard to claim 1.

Claims 10 and 11 depend from claim 9 and are, therefore, patentable over Liddy et al., Grefenstette et al., and Olsson et al., whether taken alone or in any reasonable combination for at least the reasons given with regard to claim 9.

Claim 17 depends from claim 16 and is, therefore, patentable over Liddy et al., Grefenstette et al., and Olsson et al., whether taken alone or in any reasonable combination for at least the reasons given with regard to claim 16. Moreover, claim 17 recites features similar to, yet possibly of different scope than, features recited above with respect to claim 5. Therefore, claim 17 is patentable over Liddy et al., Grefenstette et al., and Olsson et al., whether taken alone or in any reasonable combination, for at least reasons similar to the reasons given above with respect to claim 5.

Claim 19 depends from claim 18 and is, therefore, patentable over Liddy et al., Grefenstette et al., and Olsson et al., whether taken alone or in any reasonable combination for at least the reasons given with regard to claim 18.

Application No. 10/685,479  
Amendment dated March 17, 2008  
Reply to Office Action of November 16, 2007

Docket No.: BBNT-P01-199

In view of the above amendment, Applicant believes the pending application is in condition for allowance. To the extent necessary, a petition for an extension of time under 37 C.F.R. § 1.136 is hereby made.

Applicant believes no fee is due with this response other than that set forth on the enclosed fee transmittal. However, if a fee is due, please charge our Deposit Account No. 18-1945, under Order No. BBNT-P01-199 from which the undersigned is authorized to draw.

Dated: March 17, 2008

Respectfully submitted,

/Michael J. Chasan/

Michael J. Chasan

Registration No. 54,026

ROPES & GRAY LLP

One International Place

Boston, Massachusetts 02110

(617) 951-7000

(617) 951-7050 (Fax)

Attorneys/Agents For Applicants